

IBM FlashSystem in OLTP Database Environments

IBM Redbooks Solution Guide

IBM® FlashSystem storage systems deliver extraordinary business value with extreme performance, efficiency, and reliability through easy-to-deploy all-flash systems. IBM's recent acquisition of Texas Memory Systems (TMS), an IBM Company, extends IBM's leadership in flash optimized storage infrastructures. By using patented Variable Stripe RAID™ technology and 2D Flash RAID™, FlashSystem storage systems enhance system resiliency without sacrificing performance or usable capacity. For a truly self-optimized and automated storage infrastructure, you can combine FlashSystem storage systems with IBM System Storage® SAN Volume Controller virtualization and IBM System Storage Easy Tier®, intelligent data placement software. The result is further optimization of workloads, improved performance and service levels, and faster response to changing business requirements.

Figure 1 illustrates the value of the IBM FlashSystem storage infrastructure.

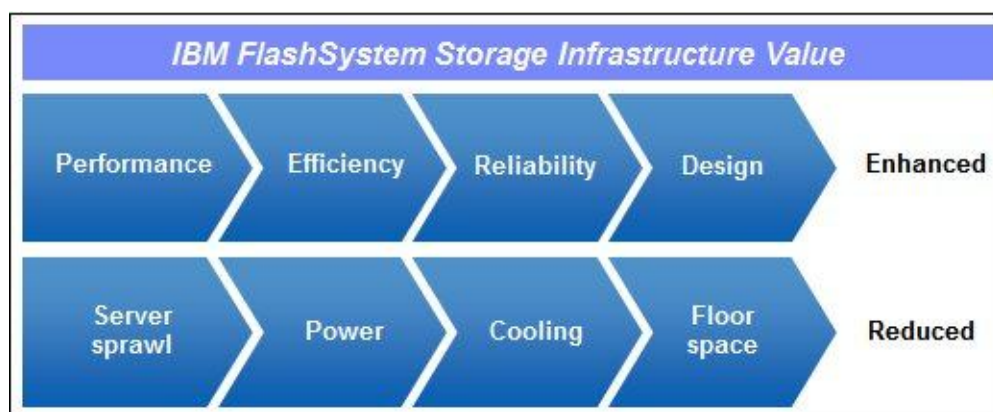


Figure 1. IBM FlashSystem storage infrastructure value

FlashSystem storage systems transform the data center environment and enhance performance and resource consolidation to gain the most from business processes and critical applications. Examples of such processes and applications include online transaction processing (OLTP), business intelligence (BI), online analytical processing (OLAP), virtual desktop infrastructures, high-performance computing, and content delivery solutions (such as cloud storage and video on demand). This solution guide focuses on IBM flash storage solutions for OLTP database environments, which are based on FlashSystem storage systems.

Did you know?

FlashSystem storage systems deliver over 500,000 read IOPS with less than 100 microseconds latency, while they provide up to 20 TB of usable 2D Flash RAID protected data storage just in 1U of rack space. FlashSystem storage systems have enterprise-level availability and reliability with no single point of failure, multiple layers of data correction, chip redundancy, and redundant hot swap components.

Business value

OLTP workloads are characterized by small, interactive transactions that generally require subsecond response times. The key performance indicator (KPI) of the transactional system is latency, because the user expects to receive the requested product information or to place an order quickly. Inability to meet these user expectations leads to customer dissatisfaction and revenue loss. IBM flash storage addresses these challenges by providing low latency, extreme performance, and efficient transaction management.

For most OLTP systems, the processor, memory, and I/O subsystem in a server are well balanced and are not considered performance bottlenecks. The major source of performance issues in OLTP environments is typically related to the storage I/O activity. The speed of traditional hard disk drive (HDD)-based storage systems does not match the processing capabilities of the servers. As a result, often a situation occurs where a powerful processor sits idle, waiting for the storage I/O requests to complete, negatively impacting user and business productivity. The negative impact on productivity extends the time to return on investments (ROI) and increases overall total cost of ownership (TCO). Therefore, storage IOPS performance and latency become strategic considerations for business. It is critical to ensure that the response time goals are met and that performance optimization is realized for other system resources (processor and memory).

In general, clients might experience the following challenges in OLTP environments:

- Failure to meet user expectations and service levels because of slow application response time
- Decreased user and business productivity
- Application and data availability concerns (slow batch processing, long backup windows, and hardware failure rates)
- Increased storage performance and capacity requirements
- Scalability constraints because of data center space, power, and cooling limits
- Increasing TCO
 - Rising data center power and cooling costs
 - Increasing software licensing fees
 - Rising server, network, and storage infrastructure management and support costs
- Longer lead time to ROI because of inefficient utilization of the existing resources

FlashSystem storage systems address the challenges in the following ways:

- Dramatically boosting the performance of existing applications and lowering cost per IOPS ratio without a need for rearchitecture
- Increasing user productivity with better response times, improving business efficiency
- Increasing data availability by using advanced system-level high availability and reliability technologies, reducing the number of solution components and shortening batch processing and backup times
- Increasing storage performance and capacity while decreasing power, cooling, and space requirements

- Reducing TCO
 - Reducing energy costs because of lower power and cooling requirements
 - Reducing the number of systems, devices, and components that are required to build the solution by increasing usage of available resources
 - Reducing software license fees because fewer systems or processors are required
 - Reducing management and support costs because of fewer components to deploy and support
- Faster ROI because of better resource usage

Solution overview

An OLTP solution with FlashSystem storage systems consists of the following components:

- Database servers (IBM System x® or IBM Power Systems™) that run data management software such as IBM DB2®, Microsoft SQL Server, or Oracle Database
- Flash storage systems (IBM FlashSystem 720 or IBM FlashSystem 820) that host the entire data set or subsets of data
- Storage area network (SAN) that is used to provide connectivity across database servers and storage systems

IBM DB2 for Linux, UNIX, and Windows is the database of choice for robust, enterprise-wide solutions that handle high-volume workloads. It is optimized to deliver industry-leading performance while lowering costs, and IBM servers that run DB2 are proven performance leaders. DB2 uses and optimizes multiple threads automatically, with no change to applications. The unique clustering design of DB2 provides near linear scalability, continuous availability, and simplified management.

IBM System x servers are developed, tested for quality, and certified by IBM, and they are backed by IBM's incomparable worldwide service and support. IBM System x servers deliver business value over the long term because of advancements in scalability, reliability, and performance. These advancements are combined with flexible configuration options, energy efficient components, and robust systems management tools.

IBM eX5 servers, such as IBM System x3850 X5, are designed for mission-critical enterprise-class workloads such as databases. By using open, industry-standard components, which are combined with IBM eX5 technologies, these systems provide leadership performance, scalability, and reliability.

Ideally suited for computing intensive workloads, IBM Power Systems deliver leadership performance and scalability in its class. An integrated approach to the design, development, and testing of each IBM POWER® server, blade, or compute node ensures the resiliency that is required for today's IT infrastructure. All Power Systems server models include innovative reliability, availability, and serviceability features that help you avoid unplanned downtime. And, with capacity on demand, hot-node add, and IBM Active Memory™ Expansion, Power Systems enterprise servers ensure that you can keep your most important applications available, even as you add capacity to handle new business demands.

IBM flash storage offerings that are used in OLTP environments include FlashSystem 720 and FlashSystem 820 purpose-built external flash storage systems. These storage systems feature one of the industry's most dense GB capacity and IOPS per 1U rack space ratios for a protected storage with low-power consumption.

Figure 2 shows FlashSystem 720, high availability SLC Flash system.



Figure 2. FlashSystem 720

Figure 3 shows FlashSystem 820, high availability eMLC Flash system.



Figure 3. FlashSystem 820

FlashSystem 720 and FlashSystem 820 have the following common features:

- 1U form factor
- Two dual-port 8 Gbps Fibre Channel controllers or dual-port QDR InfiniBand host interface modules
- Patented Variable Stripe RAID
- 2D Flash RAID: RAID-5 across flash chips and RAID-5 across flash modules
- Hot swappable flash
- Redundancy for power, data, and management
- No single point of failure

FlashSystem 720 has the following features:

- Enterprise-grade SLC flash
- 5 or 10 TB of usable storage with system-level RAID and spare flash modules
- 525,000 IOPS (4 KB reads) and 400,000 IOPS (4-KB writes)
- 25 μ s (writes) and 100 μ s (reads) latency
- 5 GBps bandwidth (reads) and 4 GBps (writes)
- Low power (350 watts)

FlashSystem 820 has the following features:

- Enterprise-grade eMLC flash
- 10 or 20 TB of usable storage with system-level RAID and spare flash modules
- 525,000 IOPS (4 KB reads) and 280,000 IOPS (4 KB writes)
- 25 μ s (writes) and 110 μ s (reads) latency
- 5 GBps bandwidth (reads) and 2.8 GBps (writes)
- Low power (300 watts)

Solution architecture

FlashSystem storage systems are connected to the heterogeneous host platforms by using Fibre Channel SAN. Host operating systems detect FlashSystem storage system as a traditional block-level device, and applications transparently interact with it. FlashSystem storage systems supports multipathing to allow redundant storage connections through SAN, including dual-port FlashSystem interface connections, dual-port host bus adapters (HBAs) on the host systems, and redundant SAN switched fabric. Figure 4 shows the FlashSystem solution architecture.

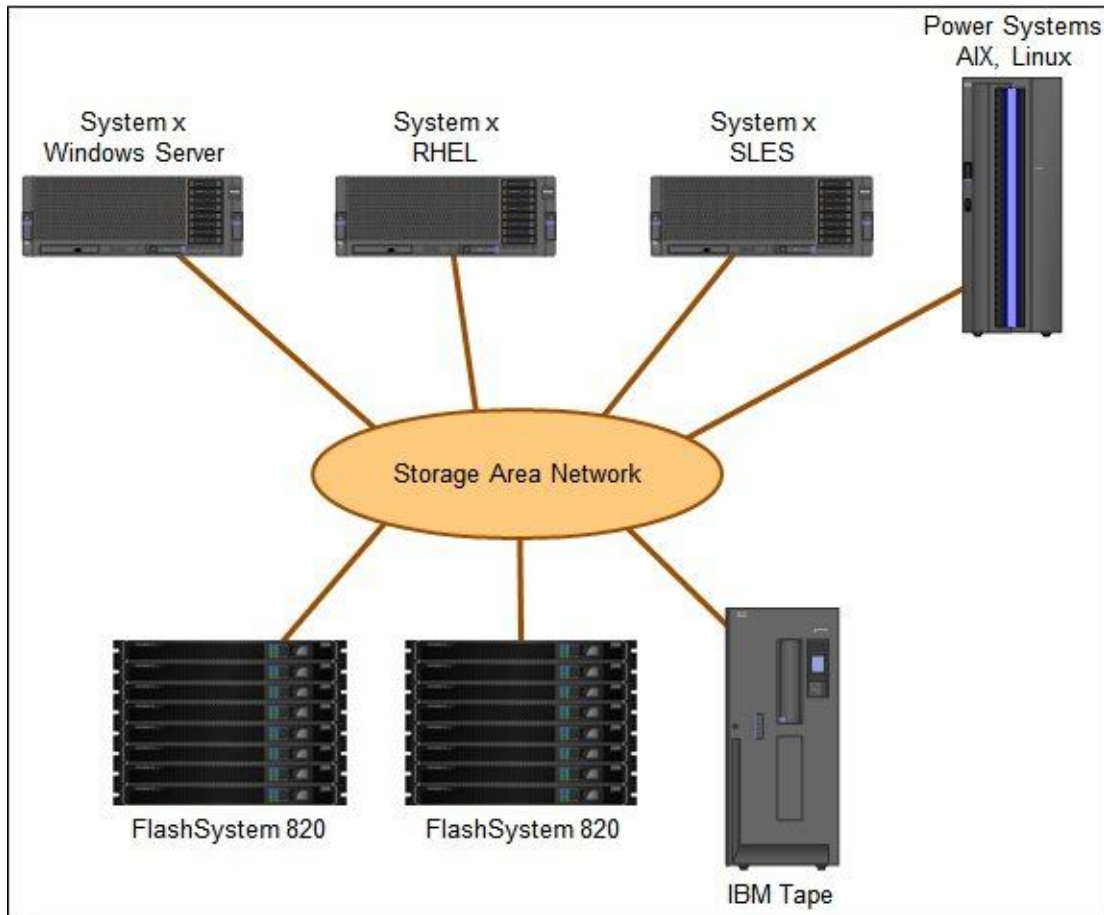


Figure 4. FlashSystem solution architecture

A single 42U rack that is filled with FlashSystem 820 storage systems can provide up to 840 TB of a protected main storage with over 20,000,000 IOPS and up to 210 GBps of throughput.

When database availability is required, you can achieve it by using either of the following ways:

- Applying database vendor-specific availability solutions, such as log shipping, replication, or database mirroring
- Using clustered systems such as Veritas Cluster Services, Microsoft Cluster Services, and Red Hat Cluster Services

Usage scenarios

FlashSystem storage systems within an OLTP environment can be used in the following ways:

- Main data storage
- Frequently accessed data storage
- Automated tired storage with IBM SAN Volume Controller

Main data storage

When FlashSystem storage system is used as a main OLTP data storage, entire database structures are placed onto FlashSystem logical volumes, as shown in Figure 5.

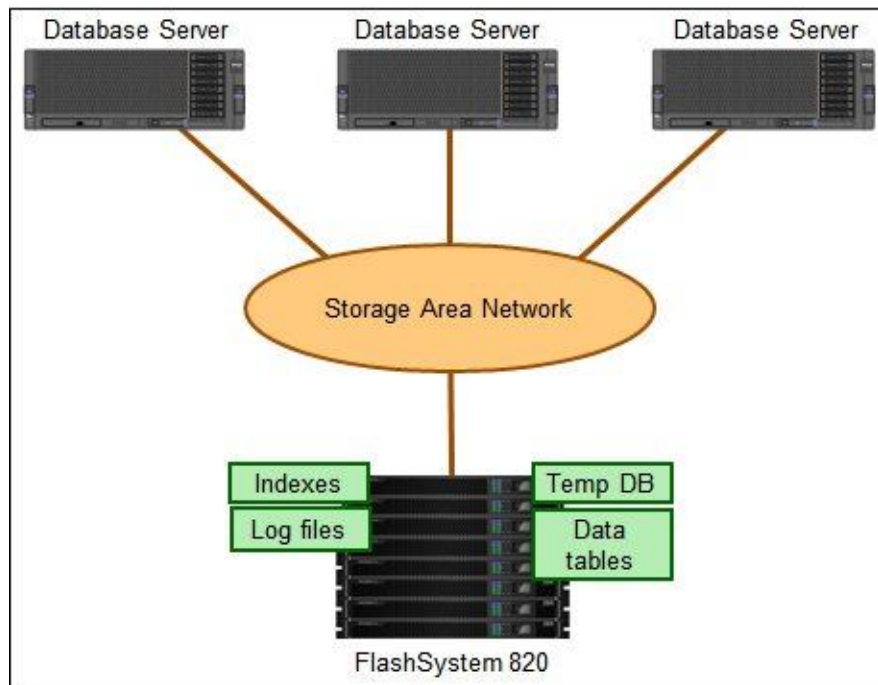


Figure 5. FlashSystem storage system as a main OLTP storage

This approach ensures maximum performance efficiency and server resource usage in heavy-loaded concurrent user access environments with frequent random access to all tables. However, if some tables are used less frequently, or rarely used, this approach might not be the most cost efficient option. Cost per GB ratio for FlashSystem storage systems is higher compared to traditional HDD-based storage systems.

Frequently accessed data storage

Sometimes, putting only part of a database structure, such as the following examples, on a high-speed FlashSystem storage system can significantly improve the performance, while keeping the storage costs optimized:

- Log files
- Temporary table space
- Frequently accessed tables
- Table partitions
- Indexes

Figure 6 shows FlashSystem storage system as a frequently accessed data storage.

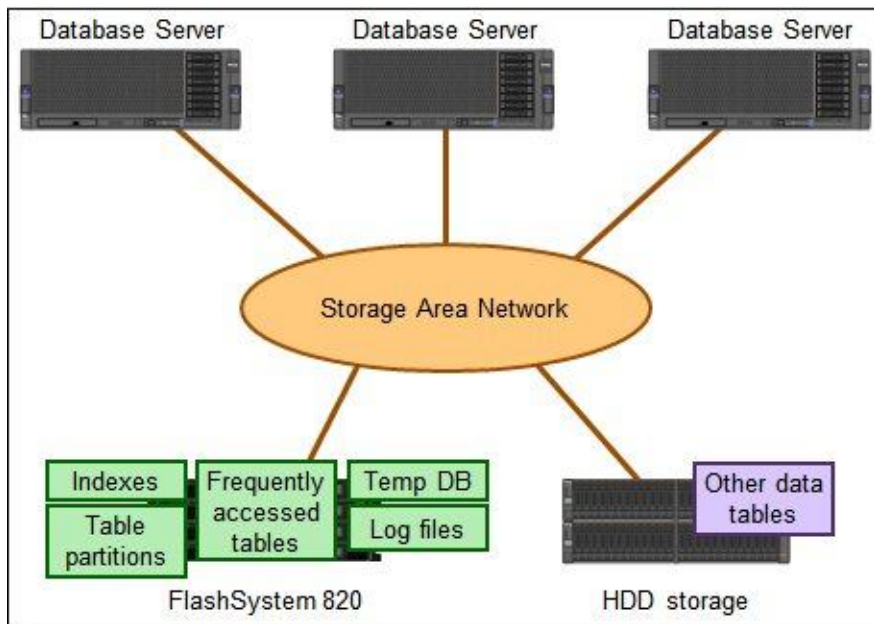


Figure 6. FlashSystem storage system as a frequently accessed OLTP storage

Although this approach can be more cost-optimized, it is manual and sometimes difficult to identify a subset of data that is frequently accessed. For these situations, automated tiered storage solutions might help.

Automated tiered storage

Automated storage tiering can help to identify most frequently used sets of data (frequently referred to as *hot data*). It provides continuous I/O monitoring and dynamic placement onto a performance optimized storage (focus on cost per IOPS). Automated storage tiering keeps infrequently used data (known as *cold data*) on a capacity optimized storage (focus on cost per GB).

FlashSystem storage systems can be combined with System Storage SAN Volume Controller and its integrated IBM System Storage Easy Tier feature to provide automated storage tiering and, therefore, dynamically optimize storage performance as illustrated in Figure 7.

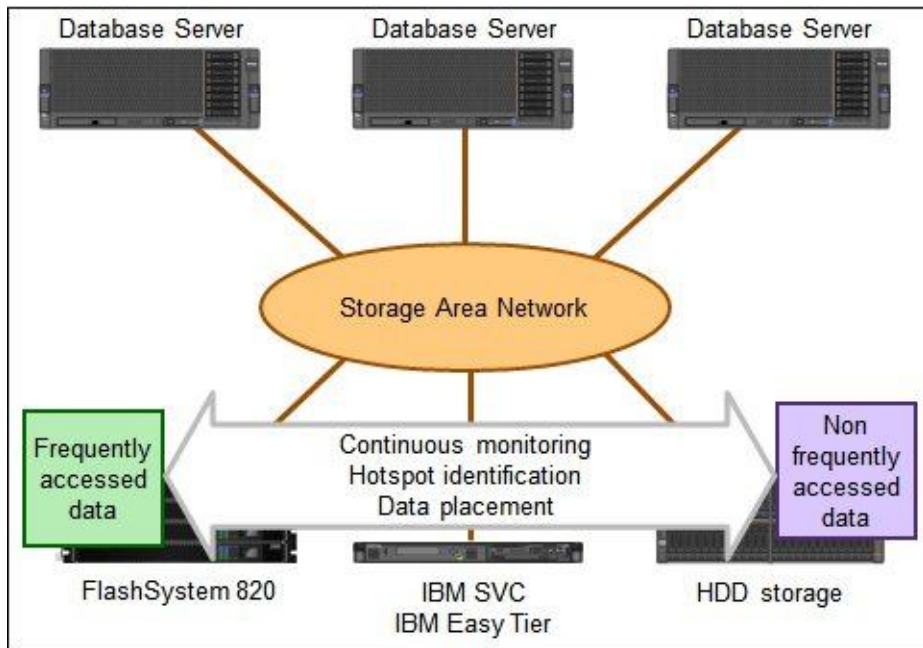


Figure 7. FlashSystem storage systems combined with SAN Volume Controller and IBM Easy Tier

IBM SAN Volume Controller provides storage virtualization through pooled management of diverse storage environments by using a simple, common interface that remains consistent regardless of storage type. SAN Volume Controller hides physical changes within the storage infrastructure to help improve availability and to streamline the task of provisioning, making dynamic management of storage assets far faster and more efficient than manual processes. To further simplify management, SAN Volume Controller integrates with IBM Tivoli® Storage Productivity Center to create a single view of physical and virtual storage resources and their relationships with physical and virtual servers.

IBM System Storage Easy Tier is a function that automatically and nondisruptively moves frequently accessed data from hard disk drives to solid-state storage, placing such data in a faster tier of storage. System Storage Easy Tier eliminates manual intervention when assigning highly active data on volumes to faster responding storage. In this dynamically tiered environment, data movement is seamless to the host application regardless of the storage tier in which the data resides.

Optionally, SAN Volume Controller supports the IBM Real-time Compression™ feature, which performs inline data compression without performance impact, helping to increase usable storage space up to five times. Unlike other approaches to compression, IBM Real-time Compression can be used with active primary data, such as production databases and email applications, which dramatically expands the range of candidate data that can benefit from compression. IBM Real-time Compression operates immediately as data is written to disk so that no space is wasted by storing uncompressed data that is awaiting post-processing.

Integration

Further extending the automated tiering scenario, IBM Tivoli Storage Productivity Center can be deployed to simplify management of heterogeneous storage infrastructure. IBM Tivoli Storage Productivity Center is an industry-leading storage resource management software that provides comprehensive visibility, control, and automation to managing heterogeneous storage infrastructures through the all-new, web-based, fully integrated management console.

Tivoli Storage Productivity Center provides storage infrastructure management capabilities such as automated system discovery, provisioning, configuration, performance monitoring, and replication for storage systems and storage networks. This tool provides storage administrators a simple way to conduct device management for multiple storage arrays and SAN fabric components from a single integrated console.

Tivoli Storage Productivity Center has the following features:

- Reduces the complexity of managing your storage environment by simplifying, centralizing, automating and optimizing storage tasks that are associated with storage systems, storage networks, replication services, and capacity management
- Provides comprehensive visibility and helps centralize the management of heterogeneous storage infrastructure from a next-generation, web-based user interface by using role-based administration and single sign-on
- Easily creates and integrates custom reports (based on IBM Cognos®) on capacity and performance
- Delivers common services for simple configuration and consistent operations across host, fabric, and storage systems
- Manages performance and connectivity from the host file system to the physical disk, including in-depth performance monitoring and analysis of SAN fabric
- Monitors, manages, and controls (zones) SAN fabric components
- Monitors and tracks the performance of SAN-attached storage devices that are compliant with SMI-S
- Manages advanced replication services (Global Mirror, Metro Mirror, and IBM FlashCopy®)

Supported platforms

FlashSystem storage systems support a wide range of operating systems (Windows Server 2003 and 2008, Linux, and AIX), hardware platforms (System x, Power Systems, and x86 servers not from IBM), HBAs, and SAN fabrics. For specific information, see the System Storage Interoperation Center (SSIC): <http://ibm.com/systems/support/storage/ssic>

Ordering information

For FlashSystem 720 and FlashSystem 820 ordering information, see the following IBM Redbooks® Product Guide:

- *IBM FlashSystem 720 and IBM FlashSystem 820*
<http://www.redbooks.ibm.com/abstracts/tips1003.html>

Related information

For more information, see the following documents:

- IBM FlashSystem family product page
<http://www.ibm.com/storage/flash>
- *IBM FlashSystem 720 and IBM FlashSystem 820 Product Guide*, TIPS-1003
<http://www.redbooks.ibm.com/abstracts/tips1003.html>
- IBM System x product page
<http://www.ibm.com/systems/x>
- IBM Power Systems product page
<http://www.ibm.com/systems/power>
- IBM SAN Volume Controller product page
<http://www.ibm.com/systems/storage/software/virtualization/svc/index.html>
- *Implementing the IBM System Storage SAN Volume Controller V6.3*, SG24-7933
<http://www.redbooks.ibm.com/abstracts/sg247933.html>

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service. IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you. This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk. IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you. Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products. This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurement may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs.

© Copyright International Business Machines Corporation 2013. All rights reserved.

Note to U.S. Government Users Restricted Rights -- Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

This document was created or updated on April 9, 2013.

Send us your comments in one of the following ways:

- Use the online **Contact us** review form found at:
ibm.com/redbooks
- Send your comments in an e-mail to:
redbook@us.ibm.com
- Mail your comments to:
IBM Corporation, International Technical Support Organization
Dept. HYTD Mail Station P099
2455 South Road
Poughkeepsie, NY 12601-5400 U.S.A.

This document is available online at <http://www.ibm.com/redbooks/abstracts/tips0973.html> .

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. These and other IBM trademarked terms are marked on their first occurrence in this information with the appropriate symbol (® or ™), indicating US registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at <http://www.ibm.com/legal/copytrade.shtml>

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:

Active Memory™
AIX®
Cognos®
DB2®
Easy Tier®
FlashCopy®
IBM®
POWER®
Power Systems™
Real-time Compression™
Redbooks®
Redbooks (logo)®
System Storage®
System x®
Tivoli®

The following terms are trademarks of other companies:

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Microsoft, Windows, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Other company, product, or service names may be trademarks or service marks of others.